Distributed Data Fusion and Maritime Domain Awareness for Harbor Protection

Simon J. Julier[a1] and Ranjeev Mittu[b2]
[a1]*ITT Advanced Engineering Systems / Naval Research Laboratory, Washington DC*
[b2]*Naval Research Laboratory, Washington DC*

Abstract. Protecting a harbor against intentional and accidental threats is extremely difficult. Harbors are not closed systems. Rather, they are focal points for the movement of people and cargo, both in land and on water. As such, threats can arise from many sources that range from the smuggling of illegal and dangerous goods to the placement of mines to damage or destroy shipping. To detect the many different types of threats, a harbor must be monitored by multiple sensing systems with different sensing modalities. To be practical, such a large sensing system must be cost effective to install, can be readily upgraded, and should be robust to sensor and communication failures. In this chapter we discuss the role that distributed data fusion can play in harbor protection. We define and discuss distributed data fusion algorithms and illustrate how they could be used in port surveillance and Maritime Domain Awareness applications.

Keywords: distributed data fusion, harbor security, maritime domain awareness, distributed operations

1. Introduction

The need for Harbor Protection is extremely important. Harbors are critical for commercial activities. However, the sheer volume of people and material moving through them means that they are not closed systems. Rather, threats can arise in many different ways from many different sources. These include the use of commercial vessels for contraband smuggling and trafficking (people and/or weapons), the potential use of commercial vessels to support other illegal activities that could lead to terrorist activities, and threats that directly impact the harbor itself (such as mining the waterways). These difficulties are exacerbated by the fact that some types of threats — such as the hijacking of commercial ships — means that an effective terrorist attack can be initiated even while the ship is far from port and there is a substantial time before the threat manifests itself [1]. Therefore, threats should be detected as far away and as early as possible before they have an opportunity to reach their destination. In the best case, detection is achieved before the threat departs from port headed towards the destination. However, detection may occur while the vessels are in transit but at a sufficient distance from the destination. Therefore, the ability to recognize, monitor, track and intercept suspect maritime vessels on a global scale is being seen as a major capability that will enable the United States and its allies to stop future global terrorist activities. This capability, known as Maritime Domain Awareness (MDA) is being pursued by many agencies in the Department of Defense (DoD).

Whatever the source of the threat, one means of identifying and responding to it is to start with an accurate Maritime Common Operational Picture (MCOP). The MCOP is formed by integrating multi-source intelligence information obtained through a worldwide network. The information may contain raw measurements that are fused with other raw measurements (Level 1 fusion) to enable the estimation of objects including their identity and kinematics. Level 1 fusion is a necessary precursor that enables Level 2 fusion, which is concerned with situation assessment and the ability to recognize activities and their relationships. Level 3 fusion concerns itself with threat assessment and ability to reason about entity intent. Generally, systems that provide the MCOP are concerned with Level 1 fusion. However, as the DoD moves towards the vision of realizing network-centric warfare operations, it is reasonable to expect that services supporting Level 2/3 fusion will be available.

---

[1] Now with the Department of Computer Science, University College London.

[2] Corresponding author: Advanced Information Technology Division Code 5580, Naval Research Laboratory, 4555 Overlook Avenue SW, Washington DC 20375, USA. Email: ranjeev.mittu@nrl.navy.mil.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
OCT 2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Distributed Data Fusion and Maritime Domain Awareness for Harbor Protection	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,4555 Overlook Avenue SW,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Aspects of Network and Information Security, NATO Advanced Studies Institute on Network Security and Intrusion Detection, Nork, Yerevan, Armenia, IOS Press, Oct 2006.

14. ABSTRACT
Protecting a harbor against intentional and accidental threats is extremely difficult. Harbors are not closed systems. Rather, they are focal points for the movement of people and cargo, both in land and on water. As such, threats can arise from many sources that range from the smuggling of illegal and dangerous goods to the placement of mines to damage or destroy shipping. To detect the many different types of threats, a harbor must be monitored by multiple sensing systems with different sensing modalities. To be practical, such a large sensing system must be cost effective to install, can be readily upgraded, and should be robust to sensor and communication failures. In this chapter we discuss the role that distributed data fusion can play in harbor protection. We define and discuss distributed data fusion algorithms and illustrate how they could be used in port surveillance and Maritime Domain Awareness applications.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

The MCOP can be formed using the *centralized* architecture shown in Figure 1~~Figure 1~~(A): all the raw sensor data is sent to a central fusion site where it is fused together. However, an alternative is to use the *distributed* data fusion (DDF) system illustrated in Figure 1~~Figure 1~~(B). Such systems can be flexible, robust and tiered. They replace the notion that the network consists of *sensors* and a *fusion center* by a set of *processing nodes*, connected to one another through communication links [2]. Each processing node can have zero or more sensing devices attached to it. There is no single central fusion center (the system state can be extracted by a "system monitor" which can be attached to any node in the network); there is no common communication facility (all communication is managed on a node-to-node basis); there is no need for global knowledge of network topology (nodes need only know the other local nodes they communicate directly with).

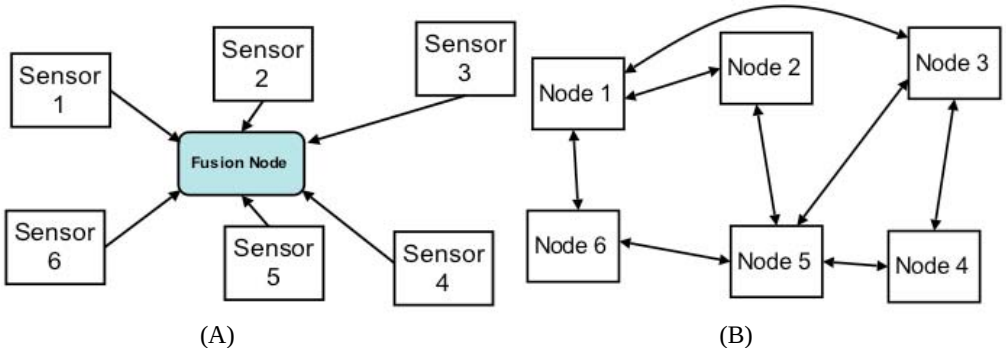

Figure 1: Centralized and distributed fusion architectures. In a centralized architecture, all the sensor data is routed to a central fusion site. In a distributed architecture, sensor data is fused throughout the network in processing nodes.

Given the potential benefits of distributed data fusion, the purpose of this chapter is to describe its basic principle of operation, discuss different types of architectures, and discuss how it can be applied to Harbor Protection. A description of DDF and the different network topologies is provided in Section 2. An application to Harbor Protection is given in Section 3. Conclusions and a summary are drawn in Section 4.

2. Distributed Data Fusion

2.1. Components of a DDF System

As explained above, a distributed data fusion (DDF) system consists of a set of processing nodes connected together through communication links. Each node possesses zero or more sensing devices. Nodes fuse data from two sources: data collected from local sensors (if available) and data distributed to it from other nodes. The communication between nodes is entirely local: a single node only knows the list of other nodes it communicates with: no single node need know the entire topology of the network.

There is no direct one-to-one correspondence between processing nodes and platforms. Figure 2 illustrates a possible configuration of processing nodes on a single platform such as an Unmanned Aerial Vehicle (UAV). The UAV possesses two types of sensors such as Forward Looking Infra-Red (FLIR), Laser Radar (LADAR) and a database. These are configured in two separate nodes: one for handling the low-level data, the other for handling the database.

There are many attractive properties to DDF systems including [2]:

- **Reliability**. The loss of a subset of nodes and/or links does not necessarily prevent the rest of the system from functioning. In a centralized system, however, the failure of a common communication manager or a centralized controller can result in an immediate catastrophic failure of the system.
- **Flexibility**. Nodes can be added or deleted by making only local changes to the network. For example, the addition of a node simply involves the establishment of links to one or more nodes in the network. In a centralized system, however, the addition of a new node can change

the topology in such a way as to require massive changes to the overall control and communications structure.

- **Bandwidth**. Nodes do not need to distribute raw sensor data. Rather, by propagating fused sensor products significant bandwidth savings can be achieved. For example, a processing node with a camera could use computer vision algorithms to process the image and identify and track targets. Therefore only the fused projects (e.g., trackID, pixel coordinates, and pixel velocity) need be distributed. In a centralized system, however, all the raw sensor data (video) would have to be transmitted to the central node to be processed.

As the example in Figure 2 shows, the capabilities of all nodes need not be the same and can vary in at least five different ways [3]:

1. **Local sensing capability**. There are many sources of intelligence information [12] (e.g., Signals Intelligence and Electronic Intelligence, to name a few) and *apriori* data (databases and other offline sources of information). However, some nodes might possess no sensors at all. Rather, they can perform the role of aggregating, forwarding and disseminating information.
2. **Signal processing**. A small unattended ground sensor, for example, might perform simple low pass filtering and use crude localization algorithms to localize a target within a fixed detection region. At its most complicated, a node might perform target class recognition using a variety of pattern recognition algorithms, constraining the results using a set of geospatial and other databases. In the most extreme case, a node might be a fusion center consisting of many analysts utilizing many types of data. It should be noted that although most DDF algorithms have been applied to Level 1 Fusion there is no difficulty, in principle, with applying these methods to Level 2 and Level 3 Fusion as well.
3. **Available bandwidth**. Different types of nodes have access to different network resources. This can depend on both the capability of the node and its current activity. As a result, the bandwidth available on different communication links can vary and signal compression schemes must be used [4].
4. **State information maintained**. Each node maintains a subset of the MCOP depending upon its sensing capabilities, purpose and security level.
5. **Roles assigned to nodes.** Depending upon hardware available, different nodes can be assigned different roles. For example, some nodes can be assigned information collection roles (significant sensing capacity; little onboard fusion), others specialize in fusion (few sensing capabilities; significant onboard fusion) and some can handle dissemination and monitoring capabilities. Furthermore, some can act as master or slave nodes.

DDF networks can be configured in a number of different network topologies, each with their own advantages or disadvantages. We now outline these.

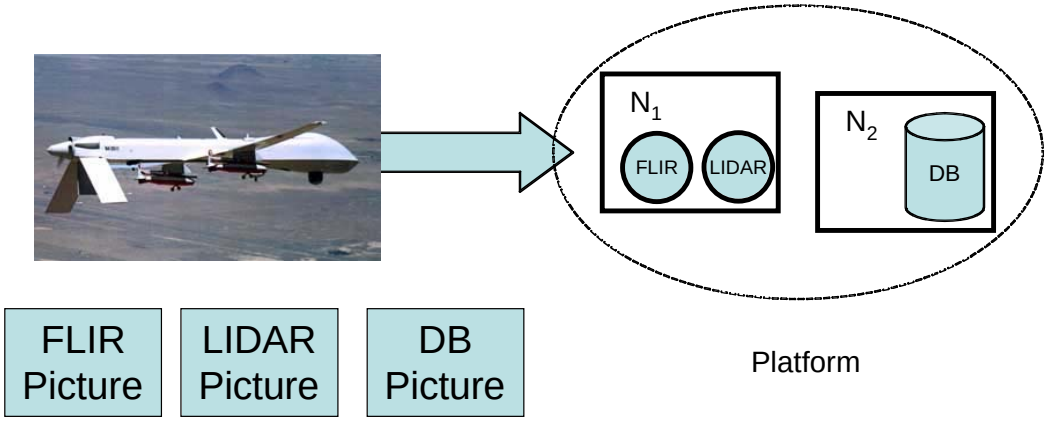

Figure 2: Configuration of processing nodes in a UAV.

2.2. Network Topologies

The different types of DDF network topologies that have been developed are illustrated in Figure 3:

- **Fully-connected**. All nodes share all their sensor information with all other nodes in a timely manner. This architecture has been deployed in the Cooperative Engagement Capability [5]. Each node has the same information and is provably optimal. However, a large number of communication links are required and the topology is brittle: if any communication link fails, the assumption that all nodes have the same state is no longer true.
- **Tree-connected**. In these networks, nodes need only communicate locally with one another and a single path exists between the nodes. Fusion algorithms have been developed which are provably optimal [6]. Furthermore, the network topology can change. However, this topology is brittle. Because there is a single path between any two nodes there is no redundancy. Adding multiple links to form redundancy leads to double-counting, as discussed below. However, the tree can reconfigure itself: if a root node is compromised, the network can be reconfigured and a new node replaced.
- **Hierarchical**. In these networks different nodes are assigned different roles. There is a central node and all data flows there through a set of intermediate nodes [7]. The intermediate nodes can, in effect, be considered a type of signal compression (for example, raw imagery data is processed to give a track of a target). However, this is a tree connected topology (hence, there still exists a central point of failure). Furthermore, the specification of roles means that there can exist a single point of failure if, for example, the master node is compromised.
- **Adhoc**. These have no special global topology. It is possible for loops and cycles to exist, leading to flexible communication architectures and redundancy. However, with general topologies no optimal local fusion algorithm can be developed.

A significant problem with distributed data fusion is there is the potential risk for *double counting*.

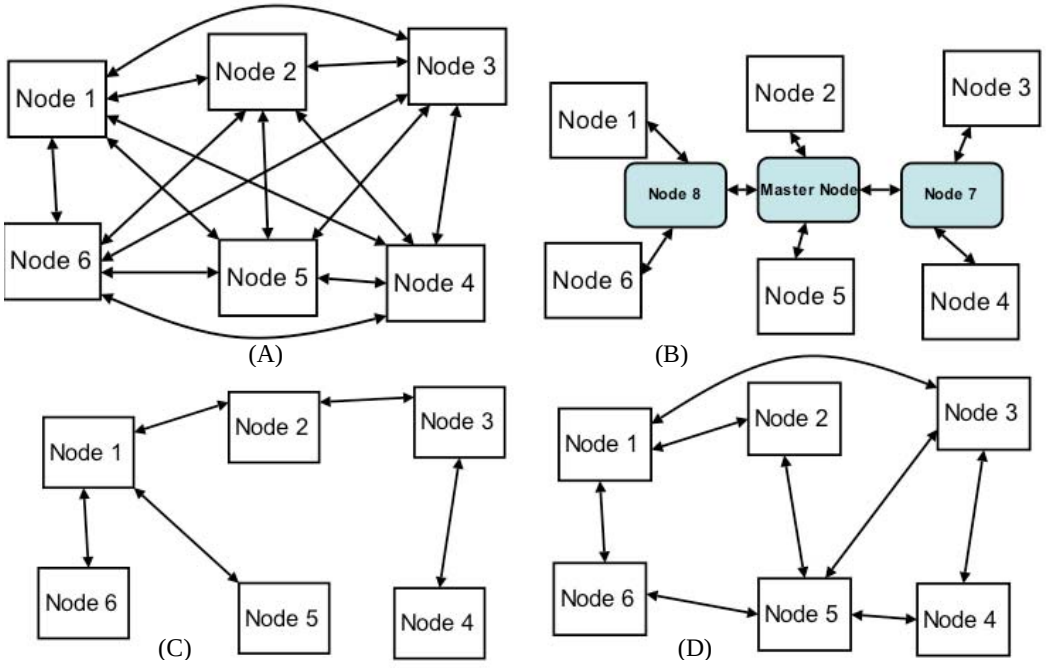

Figure 3: Various network topologies which can be used in distributed data fusion networks. (A) fully-connected; (B) tree-connected; (C) hierarchical; (D) adhoc.

2.3. Double Counting and Solutions

One of the most serious problems which may arise in a DDF network is the effect of redundant information. Specifically, pieces of information from multiple sources cannot be combined within most

filtering frameworks unless they are independent or have a known degree of correlation (i.e., known cross covariances). The effect of redundant information can be seen in the following scenario, sometimes referred to as “rumor propagation” or the “whispering in the hall” problem:

1. A node incorporating a sonar sensor detects a weak track that might be caused by an underwater threat. A hypothesis is generated that a threat exists and is propagated into the network. This information can be synopsized, augmented, or otherwise transformed as it is relayed through a sequence of nodes.
2. A threat database node receives this information and notes that a threat might be present. There are many possible interpretations of this data, but the possibility of a threat (e.g., diver) is deemed to be of such tactical importance that it warrants the transmission of a low confidence hypothesis. Again, the information can be transformed as it is relayed through a sequence of nodes.
3. The sonar sensing node receives the low confidence hypothesis that a diver threat exists. A check of available sensor data shows a feature that is consistent with the hypothesis. Because the node is unaware that the hypothesis was based on exactly the same sensor evidence, it assumes that the feature that it observes is an independent confirmation of the hypothesis. The node then transmits high confidence information that the feature represents the threat
4. The threat database node receives information from the sonar sensing node that a diver threat has been identified with high confidence. The threat database node regards this as confirmation of its early hypothesis and calls for an aggressive response to the situation.

This problem cannot be solved using optimal data fusion algorithms [8]. However several classes of suboptimal data fusion algorithms have been developed which can be used to overcome these difficulties [9].

We now illustrate how distributed DDF can be applied to MDA for Harbor Security.

3. Examples

Harbor surveillance and protection is critical in the defense of shore based assets. Therefore, threats should be detected as far away and as early as possible before they have an opportunity to reach their destination [1]. There may be many factors that help in detecting possible threats while they are in transit, and one example may be an unusual change in the vessels normal path or in its emissions pattern. Detecting the change in vessel behavior is part of the remit of MDA.

MDA is a complex problem which spans numerous dimensions, including the sharing and analysis of complex, and sometimes incomplete, data to identify vessels of interest and inferring vessel intent based on that analysis. It is a process that requires the coordination and sharing of data / information between all the entities (i.e., nodes), both mobile and fixed, that take part in the global war on terror. The process must be resilient to errors, and must be done as efficiently as possible to mitigate threats in a timely manner worldwide, but particularly in ports and harbors which represent the last stages of defense. As explained above, the concept of a node can be an extremely broad one, and can span from low-level sensing devices to entire fusion centers. We now provide two examples which illustrate the flexibility of the concept. We also discuss how DDF can be used to develop multi-tiered systems.

3.1. Distributed, Mobile Sensor Nodes

The Automated Identification System (AIS) can be used to track a vessel. However, it is a signal that can be impacted by the effects of the environment such as weather phenomena. The same applies to emissions. Because effects such as these are dynamic over time, stationary nodes may not be able to receive and process such signals all the time, leading to possible gaps in collection. One means of overcoming this difficulty is to augment fixed nodes by a set of mobile nodes that can dynamically adjust and reallocate themselves to specific areas of interest. Furthermore, they are appealing due to their lower cost of operation. For example, the use of UAVs can play a significant role to support MDA due to their flexibility in reallocation.

UAVs can be used to perform many tasks (such as information gathering and target search) in many types of environments (dynamic, uncertain, dense) with both known and unknown targets and threats. Because of the dynamic nature of the environment, fixed guaranteed communication topologies between UAVs might not exist.

The use of DDF for tracking and fusion with UAVs has already been demonstrated for picture complication and distributed tracking both in simulation and in real hardware implementations [10]. It has already been shown to be a robust and scalable solution and is highly likely to convey similar benefits for MDA.

3.2. Shore-based Data Fusion and Analysis Centers

The generalization of a node can extend to the scale of shore-based data fusion and analysis centers. These centers collect and fuse multi-source intelligence data, and coordinate the distribution of the resulting MCOP to other centers. There are numerous collection and analysis centers worldwide responsible for collecting, analyzing and disseminating information on potential maritime threats to other regional centers or to tactical military centers. The tactical military centers also manage tactical systems and data links to produce an MCOP for dissemination to operational units and command elements throughout their areas of responsibility.

These centers process data from many sources. One source of information is from other shore-based analysis centers. Because track and other data are shared frequently between systems, and because the intervening signal processing and transformation steps are extremely complicated, it is impossible to calculate the degree of common information between the two or more sites. Robust decentralized data fusion techniques, however, provides a formal mechanism for describing how the interaction between the different centers can occur.

3.3. Tiered Systems

The previous two examples have shown how DDF can be used to support fusion between piers – whether they are UAVs or entire fusion centers. However, a key enabler of a sustainable military force is the notion of a tiered system [11], and is likely one of the more complex decentralized data collection, information processing and sharing systems. Conceptually, a tiered system is an integrated, multi-tier intelligence system encompassing space and air-based sensors linked to close-in and intrusive lower tiers (Figure 4). The lower tiers are not only the critical source of intelligence; they can also serve as a key cueing device for other sensors. Given the diversity of the assets, and the fact that information will need to be shared across the horizontal and vertical planes, and the environments in which the components of a tiered system will likely operate, it is not practical to envision a single data fusion architecture. Given the diversity of the information and processing nodes that may be participating in the network, it is more reasonable to expect that a given "system of systems" should be flexible, and employ a mix of architectures most suited for the operational environment, while maintaining interoperability across the spectrum.

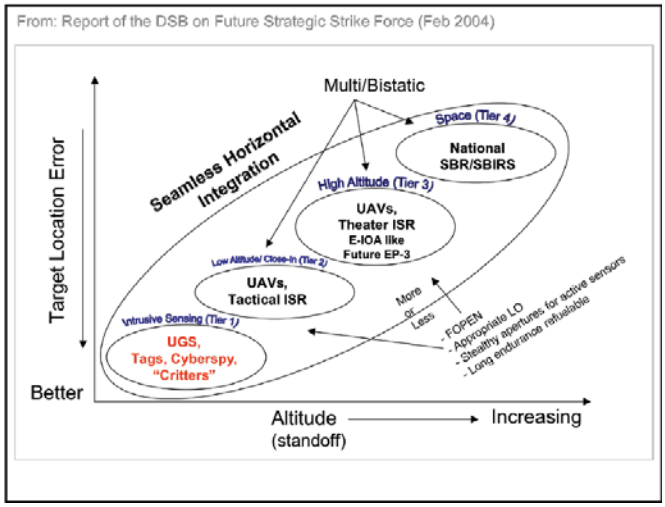

Figure 4: Tiered Systems Concept

Future maritime scenarios may involve a collection of components from a tiered-system hierarchy, with both internal communications between the tiered-system components, as well as external communications with shore-based data fusion and analysis centers. Due to the different types of information flowing within the networks ranging from raw measurements to partial estimates, coupled with decentralized system control due to the size and complexity of such an aggregate system, it is likely that a decentralized approach will provide a robust, scalable and flexible approach to data fusion.

4. Summary and Conclusions

Protecting a harbor against intentional and accidental threats is extremely difficult. Because they are focal points for the movement of people and cargo both in land and on water they are open systems. One means of identifying threats is to develop an accurate MCOP. However, to develop an accurate and consistent MCOP requires the integration of data / information from multiple sensing systems with different sensing modalities. To be practical, such a large sensing system must be cost effective to install, can be readily upgraded, and should be robust to sensor and communication failures. In this chapter we have argued that distributed data fusion networks exhibit many of these properties. We have described the properties of these networks and we have argued that algorithms to overcome double counting – one of the most important problems of such networks – can be resolved using suboptimal fusion algorithms. We have briefly described examples of how DDF can be used to coordinate teams of UAVs, fuse data between command centers, and may offer capabilities to achieve the vision of tiered systems. Given these advantages, we believe that DDF is a valuable fusion paradigm which can be used in harbor protection and MDA.

Acknowledgements

This work was supported in part by the Office of Naval Research.

5. References

[1] E. Shahbazian, M. J. DeWeert and G. Rogova, "Findings of the NATO Workshop on Data Fusion Technologies for Harbour Protection", In *Proceedings of the Photonics for Port and Harbor Security II Conference, SPIE Defence and Security Symposium*, vol. 6204, Orlando, FL, USA, 2006.

[2] H.F. Durrant-Whyte and M. Stevens, "Data Fusion in Decentralised Networks," XXX, University of Sydney, Australia, 2001.

[3] S. J. Julier, J. K. Uhlmann, J. Walters, R. Mittu and K. Palaniappan, "The Challenge of Scalable and Distributed Fusion of Disparate Sources of Information," In *Proceedings of the Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications Conference, SPIE Defense and Security Symposium*, vol. 6242, Orlando, FL, USA, 2006.

[4] D. Nicholson and V. Leung, "Managing a Distributed Data Fusion Network," In *Proceedings of the Signal Processing, Sensor Fusion and Target Recognition XIII Conference, SPIE Defence and Security Symposium*, vol. 5429, Orlando, FL, USA, 2004.

[5] "The Cooperative Engagement Capability," *Johns Hopkins APL Technical Digest*, vol. 16, no. 4, pp. 337-396, 1995.

[6] S. Grime and H. F. Durrant-Whyte, "Data Fusion in Decentralized Sensor Fusion Networks," *Control Engineering Practice*, vol. 5, no. 2, pp. 849-863, 1994.

[7] L. Y. Pao, "Distributed Multisensor Fusion," In *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Scottsdale, AZ, August 1994, pp. 82-91.

[8] S. W. Utete, "Network Management in Decentralised Sensing Systems," DPhil Thesis, Robotics Research Group, Oxford, 1995.

[9] S. J. Julier and J. K. Uhlmann, "General Decentralized Data Fusion with Covariance Intersection (CI)," In The Handbook for Data Fusion, eds. D. A. Hall and J. Llinas, Chapter 12, pp. 12-1-12-18, 2001.

[10] E. W. Nettleton, H. F. Durrant-Whyte, P. W. Gibbens, A. H. Goktogan, "Multiple Platform Localisation and Map Building", In the *Proceedings of Sensor Fusion and Decentralized Control in Robotic Systems III Conference, SPIE Photonics East*, vol. 4196, Boston, MA, USA, 2000.

[11] J. Dalburg et.al., "Developing a Viable Approach for Effective Tiered Systems," NRL Memorandum Report 1001-07-9024.

[12] [Web address] http://www.fbi.gov/intelligence/di_ints.htm, last accessed 21 March 2007